Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary information

CONSTANT PH MOLECULAR DYNAMICS OF PCV2 CAPSID PROTEIN REVEALS A MECHANISM FOR CAPSID ASSEMBLY

Elvira Tarasova*, Noriaki Okimoto, Shanshan Feng, Dmitry Nerukh, Reza Khayat*, Makoto Taiji

CONVERGENCE

Conformational landscape sampled by the capsid protein. Two-dimensional distribution of parameters SASA vs. Rg

Supplementary Figure 1. Conformation landscape sampled by CP* during simulation at pH 3. Free energy surfaces as function of parameters SASA of all residues vs. Rg of all residues from the 1000ns of the simulation. Analysis was performed for all residues in the subunit including flexible N-terminal region -amino acids 1-231. The free energy of the global minimum was set to 0 kT.

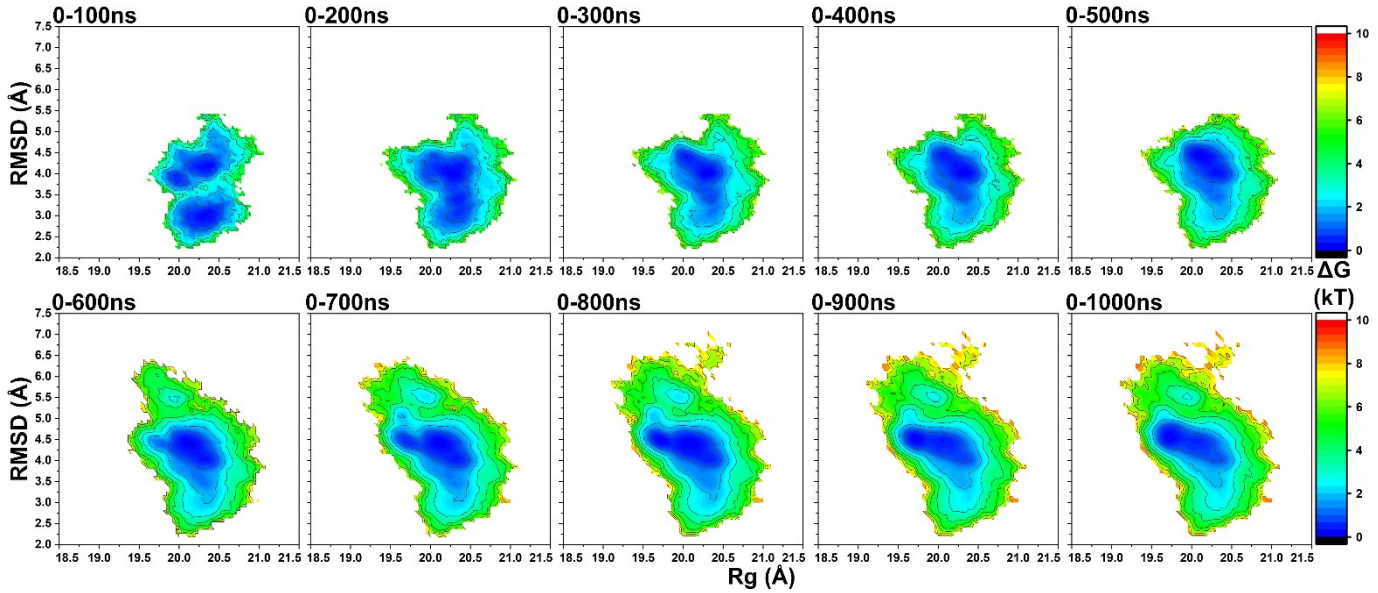
Supplementary Figure 2. Conformation landscape sampled by CP* during simulation at pH 5. Free energy surfaces as function of parameters SASA of all residues vs. Rg of all residues from the 1000ns of the simulation. Analysis was performed for all residues in the subunit including flexible N-terminal region - amino acids 1-231. The free energy of the global minimum was set to 0 kT.

Supplementary Figure 3. Conformation landscape sampled by CP* protein during simulation at pH 7. Free energy surfaces as function of parameters SASA of all residues vs. Rg of all residues from the 1000ns of the simulation. Analysis was performed for all residues in the subunit including flexible N-terminal region -amino acids 1-231. The free energy of the global minimum was set to 0 kT.

Conformational landscape sampled by the capsid protein. Two-dimensional distribution of parameters RMSD vs. Rg

a

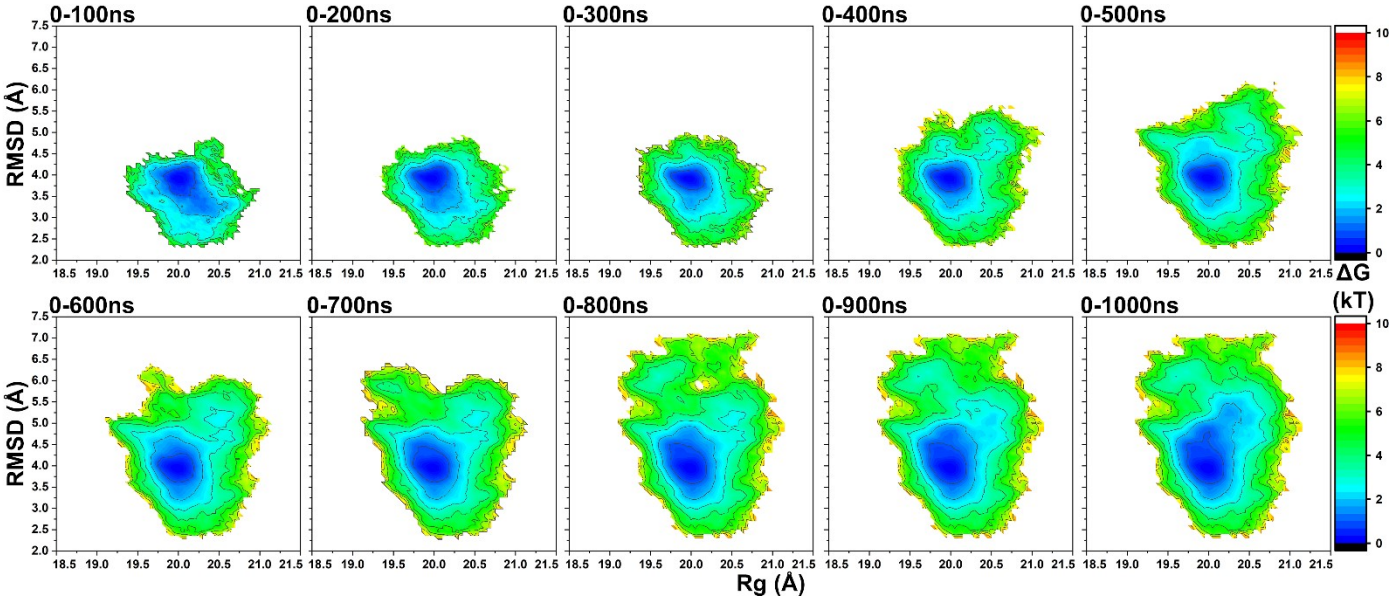
b



Supplementary Figure 4. Conformation landscape sampled by CP* during 1000ns of simulation at pH 3. Free energy surfaces as function of parameters a) RMSD of all Cα atoms vs. Rg of all residues; b) RMSD of 38-231 Cα atoms vs. Rg of 38-231 residues from the simulation. RMSD was calculated with respect to the reference X-ray structure. The free energy of the global minimum was set to 0 kT.

a

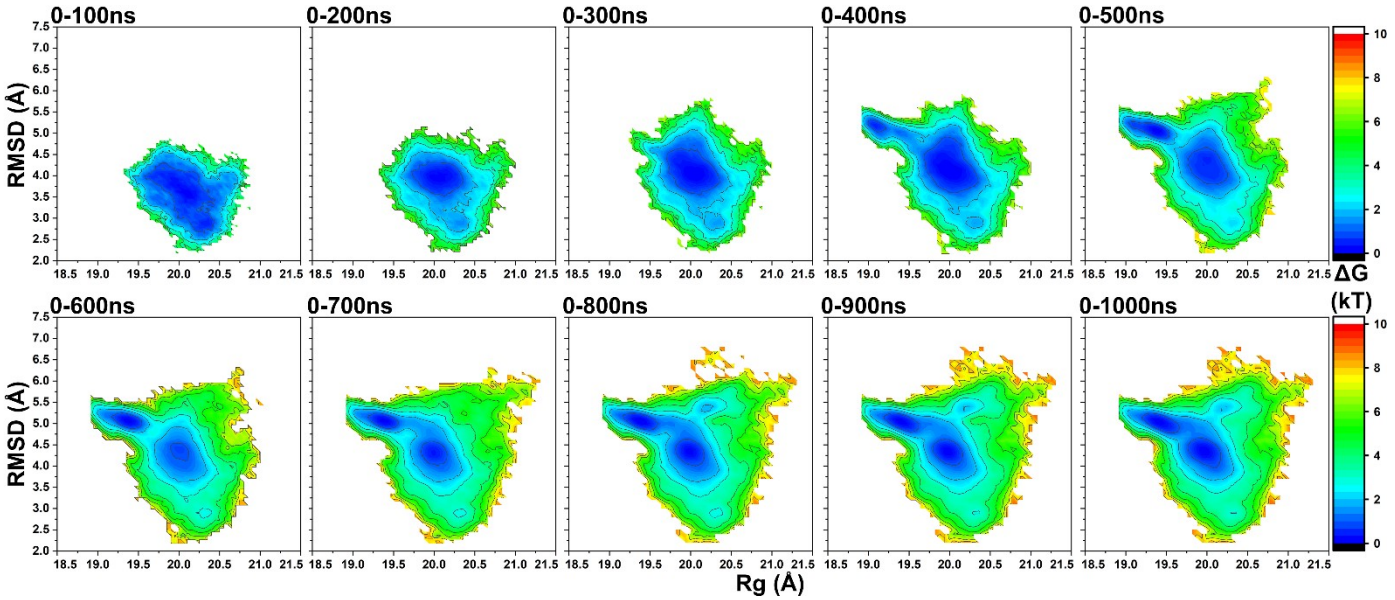
b



Supplementary Figure 5. Conformation landscape sampled by CP* during 1000ns of simulation at pH 5. Free energy surfaces as function of parameters a) RMSD of all Cα atoms vs. Rg of all residues; b) RMSD of 38-231 Cα atoms vs. Rg of 38-231 residues. RMSD was calculated with respect to the reference X-ray structure. The free energy of the global minimum was set to 0 kT.

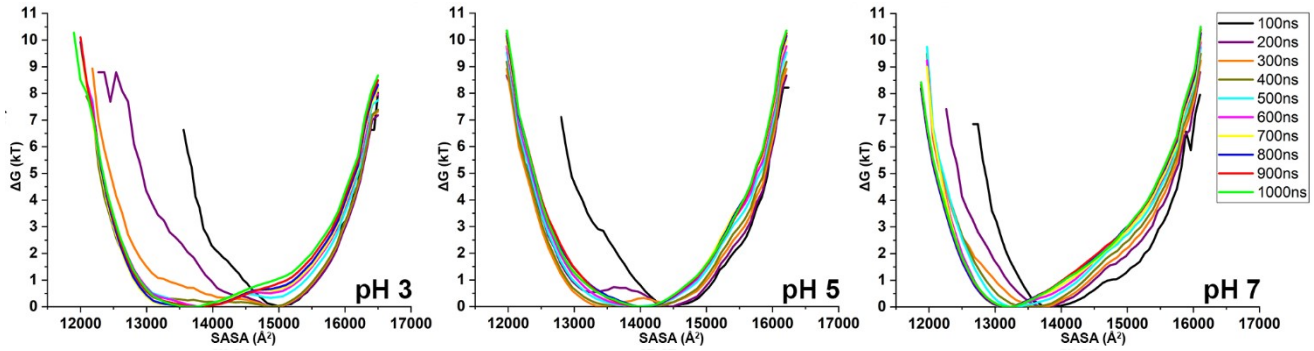
a

b

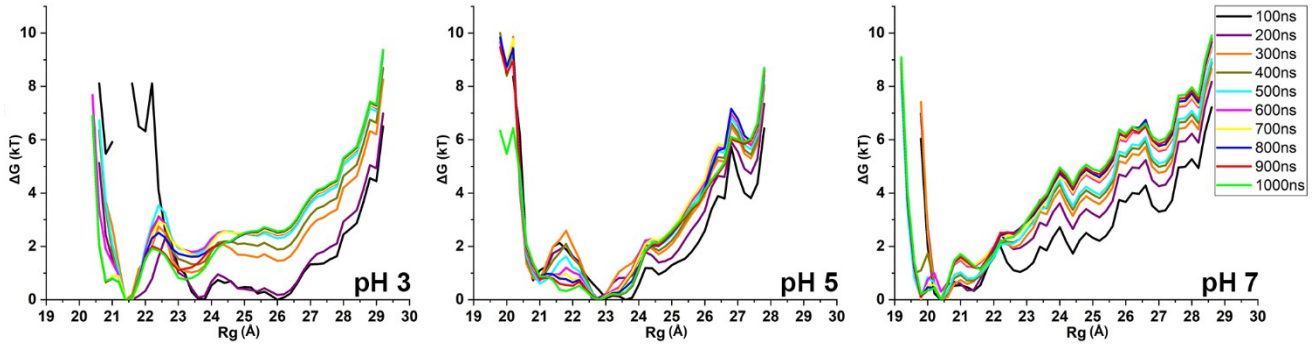


Supplementary Figure 6. Conformation landscape sampled by CP* during 1000ns of simulation at pH 7. Free energy surfaces as function of parameters a) RMSD of all Cα atoms vs. Rg of all residues; b) RMSD of 38-231 Cα atoms vs. Rg of 38-231 residues from the simulation. RMSD was calculated with respect to the reference X-ray structure. The free energy of the global minimum was set to 0 kT.

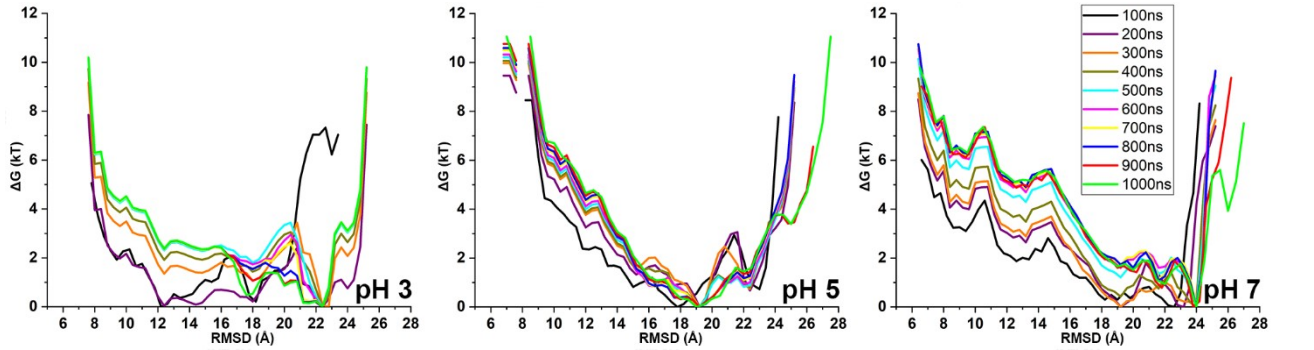
Conformational landscape sampled by the capsid protein. One-dimensional distribution of parameters SASA, RMSD, and Rg



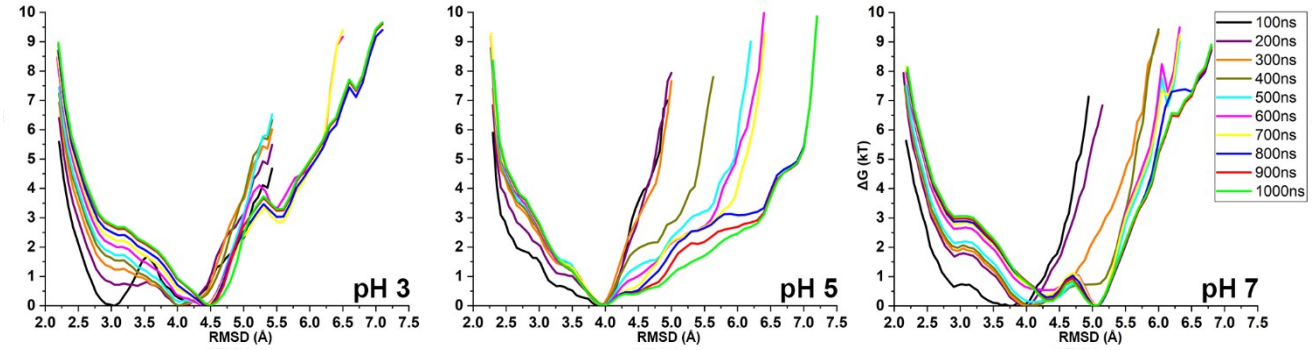
Supplementary Figure 7. Free energy diagram as a function of parameter SASA of all residues at pH 3, 5, and 7. Analysis was performed for all residues in the subunit including flexible N-terminal region -amino acids 1-231. The free energy of the global minimum was set to 0 kT.



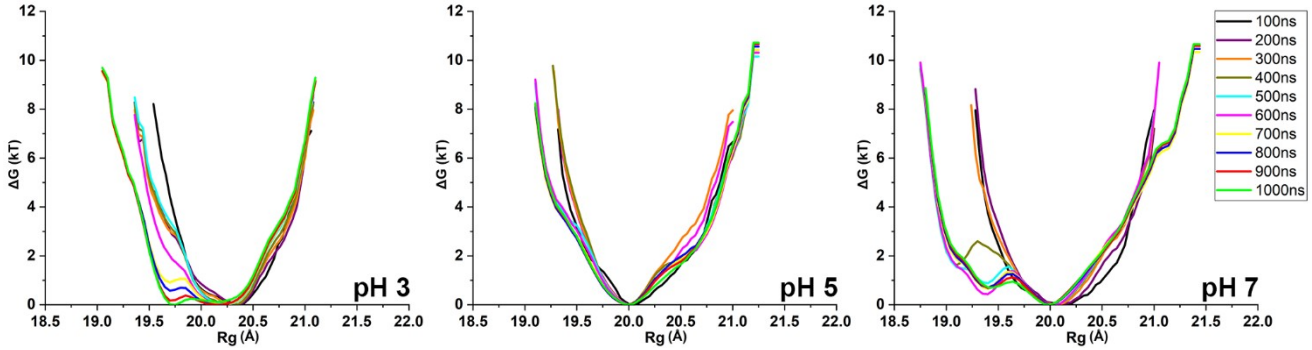
Supplementary Figure 8. Free energy diagram as a function of parameter Rg of all residues at pH 3, 5, and 7. Analysis was performed for all residues in the subunit including flexible N-terminal region - amino acids 1-231. The free energy of the global minimum was set to 0 kT.



Supplementary Figure 9. Free energy diagram as a function of parameter RMSD of all residues at pH 3, 5, and 7. Analysis was performed for all residues in the subunit including flexible N-terminal region -amino acids 1-231. *T*he free energy of the global minimum was set to 0 kT.



Supplementary Figure 10. Free energy diagram as a function of parameter RMSD of 38-231 Cα atoms at pH 3, 5, and 7. RMSD was calculated with respect to the reference X-ray structure. *T*he free energy of the global minimum was set to 0 kT.



Supplementary Figure 11. *Free e*nergy *diagram* as a function of parameter Rg of 38-231 residues at pH 3, 5, and 7. *T*he free energy of the global minimum was set to 0 kT.

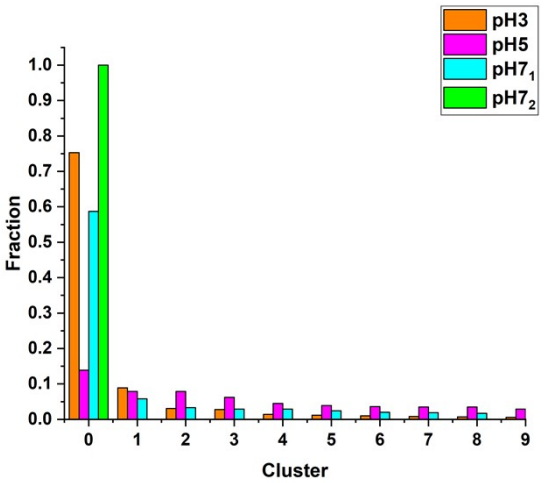
a **b** **c**

Supplementary Figure 12. Cumulatively calculated pKa for all Asp/Glu/His residues from simulations at pH 3 (a), pH5 (b), pH 7 (c) as a function of simulation time.

Supplementary Figure 13. Exchange plots for each replica. Dependence of exchanges between pH values from simulation time for individual replicas. Acceptance ratio is 21%.

REPRESENTATIVE STRUCTURES

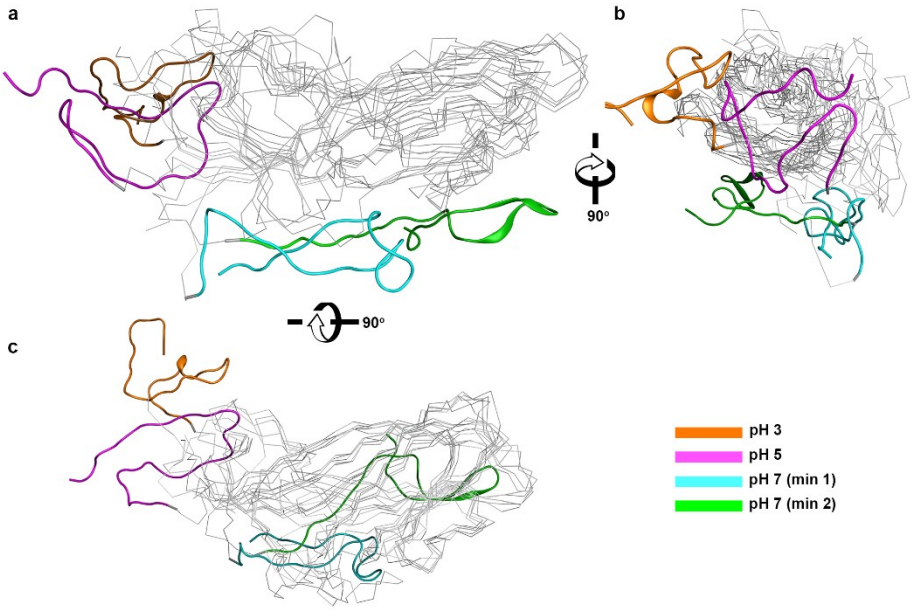
Cluster analysis suggests that the CP is more conformationally variable at pH 5



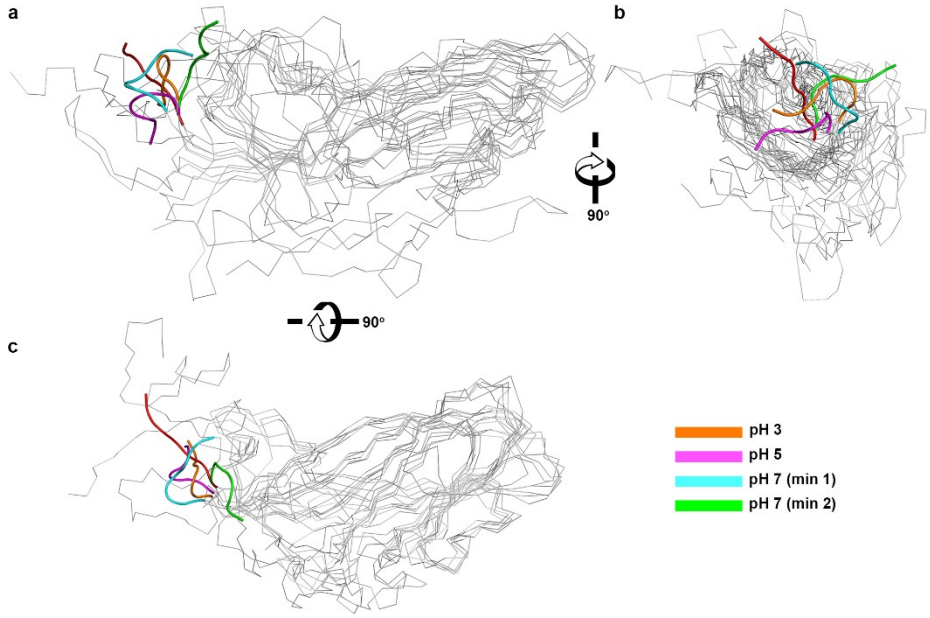
Supplementary Figure 14. Bar chart showing fraction of extracted structures per cluster for the top 10 clusters. Number of clusters starting from 0.

At pH3, the first cluster describes 75% of the conformations, the second cluster describes 10% of the conformations, and third cluster describes 4% of the conformations (Supplementary Fig. 14). At pH 5, the first cluster describes 14% of the conformations, the second and the third clusters each describe slightly less than 8% of the conformations. For the first minimum at pH 7, the first cluster describes 59% of the conformations, the second cluster describes 6% of the conformations, and the third cluster describes 3% of the conformations (Supplementary Fig. 14). The predominant majority of conformations at pH 3 and 7 are described by the first cluster, suggesting that the CP adopts a single conformation within each of these pH values. However, comparable distribution of conformations in the first three clusters of pH 5 suggests that the CP may be more conformationally variable at this pH.

Comparison of the representative structures at different pH values



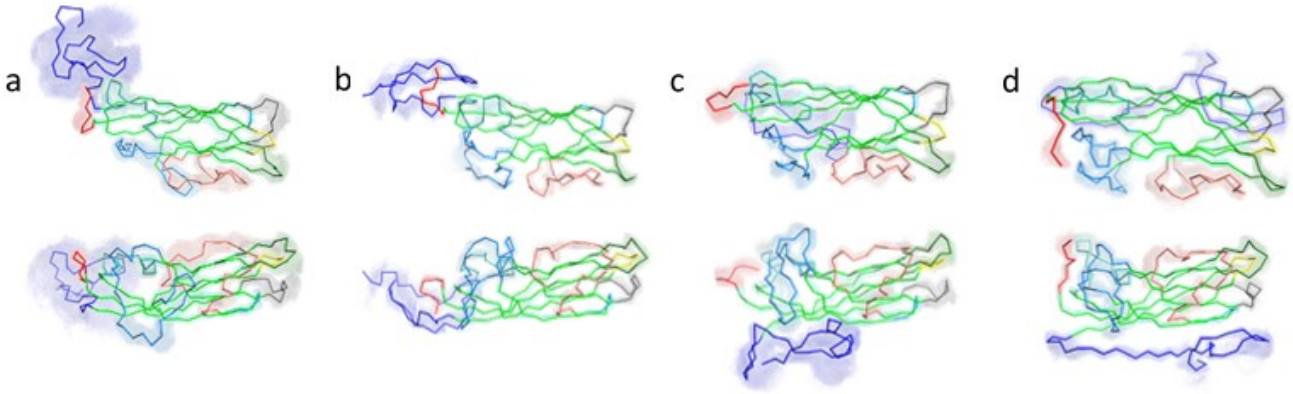
Supplementary Figure 15. Alignment of representative structures from most populated clusters at pH 3, 5, 7 on the crystal structure: N-terminals are colored, while the rest of the structure is in grey.



Supplementary Figure 16. Alignment of representative structures from most populated clusters at pH 3, 5, 7 on the crystal structure: C-terminals are colored, while the rest of the structure is in grey.

Conformational variability

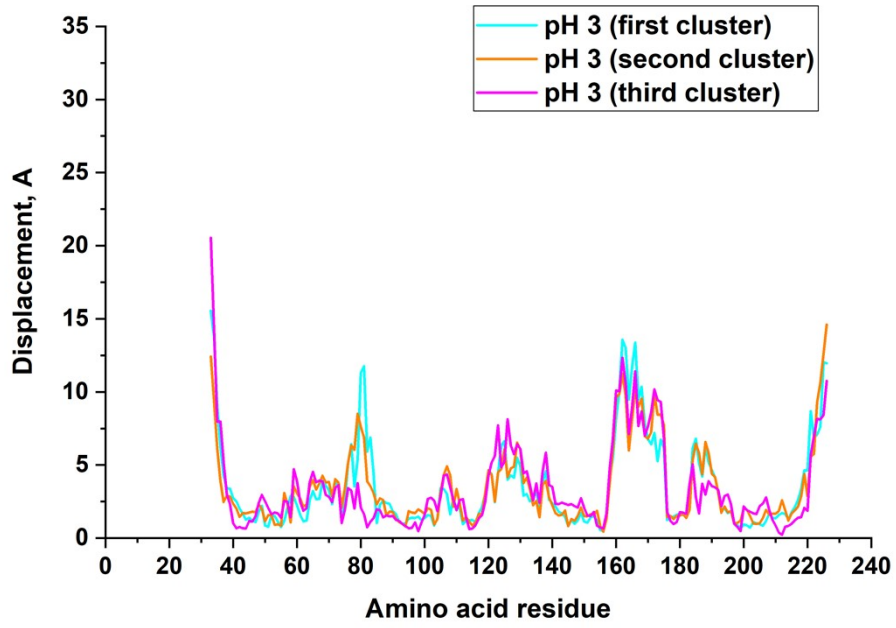
To visualize the conformational variability within each cluster, we generated a composite image of all cluster members by overlaying each member onto the centroid (Figure S17). While qualitative, these images communicate both the magnitude and direction of variability. This is advantageous over traditional quantitative methods like root mean square fluctuation that capture the magnitude of the variation but disregard the direction of the variation. As anticipated, the N-termini exhibit conformational variability; however, the N-termini of each cluster locate to comparable positions with respect to the remaining portion of the CP. This is unexpected as the N-termini were not included in the cluster analysis. The first cluster of pH 3 contains 76537 frames (Figure S17a). The N-terminus occupies a large space that is opposed to a single face of CP. The first cluster of pH 5 contains 5724 frames. The N-terminus occupies a smaller space, as compared to pH 3, and its position and orientation is consistent with respect to the remaining portion of CP (Figure S17b). The first cluster of pH 7 (minimum1) contains 19513 frames. The N-terminus occupies a larger space; however, this envelope has the same general shape and occupies comparable positions within the cluster – along inner side of the CP. The first cluster of pH 7 (minimum 2) contains 25076 structures. The conformational variability of the N-terminus is limited, and locates to the face of CP that defines the inner surface of the capsid.



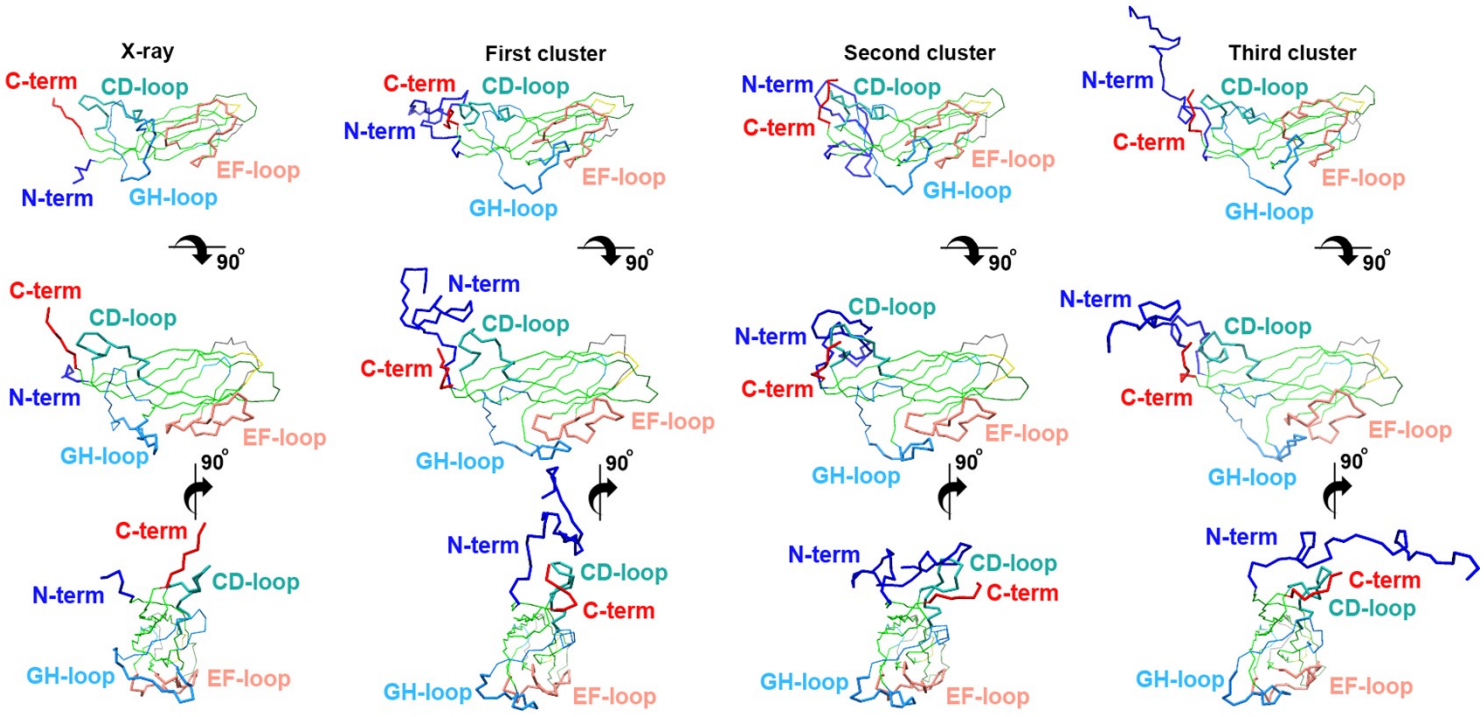
Supplementary Figure 17. Conformational variability of the first clusters at pH 3 (a), pH 5 (b), pH 7_1 (c), and pH 7_2 (d). Clustering was performed using amino acids 38-231; however, images are generated using all the amino acids. These images show the envelope (thin lines) of all frames in the cluster. Thin envelopes demonstrate small rmsd between frames whereas wide envelopes demonstrate large rmsd.

Comparison of the centroids from first three clusters at each pH

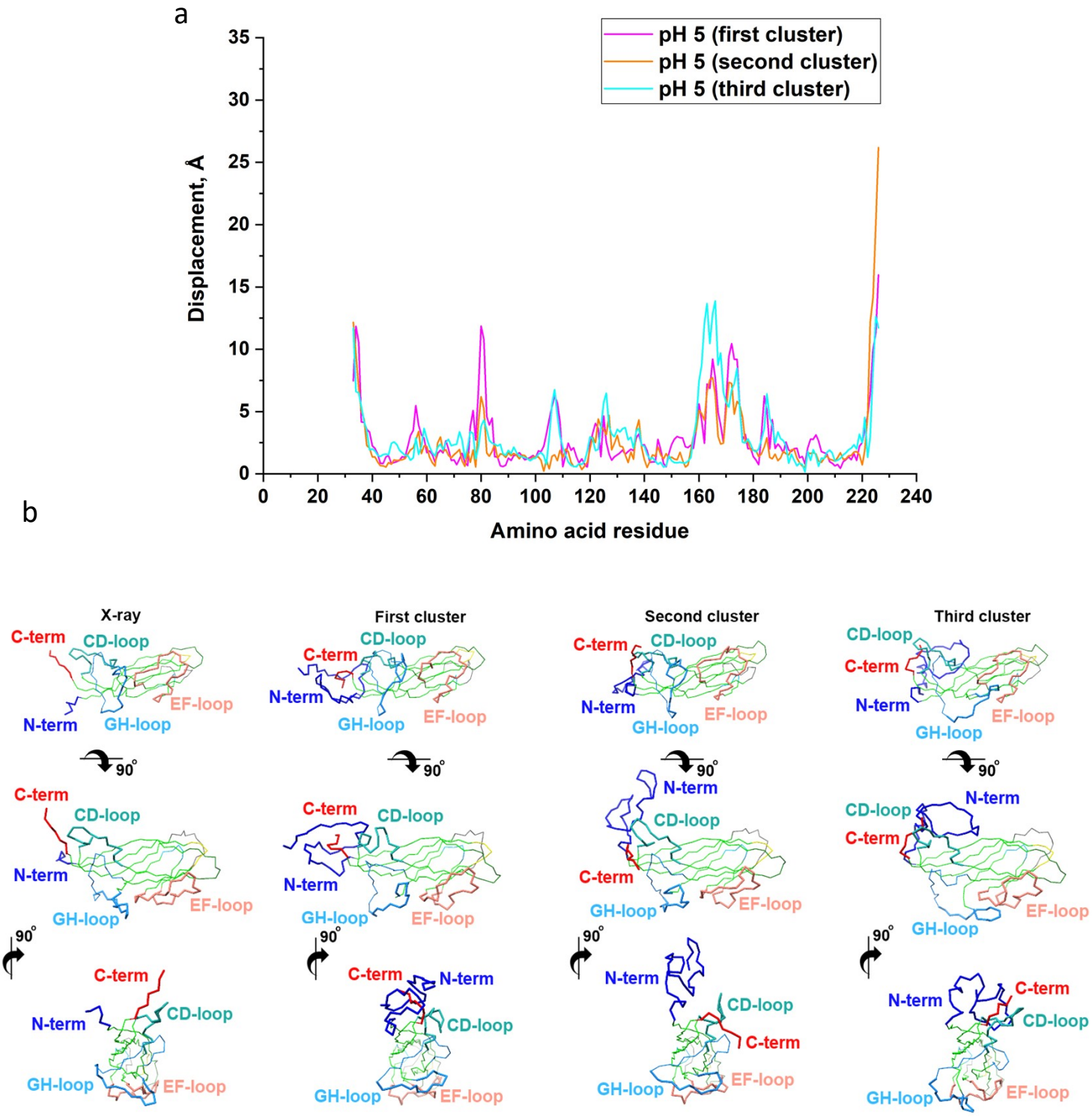
a



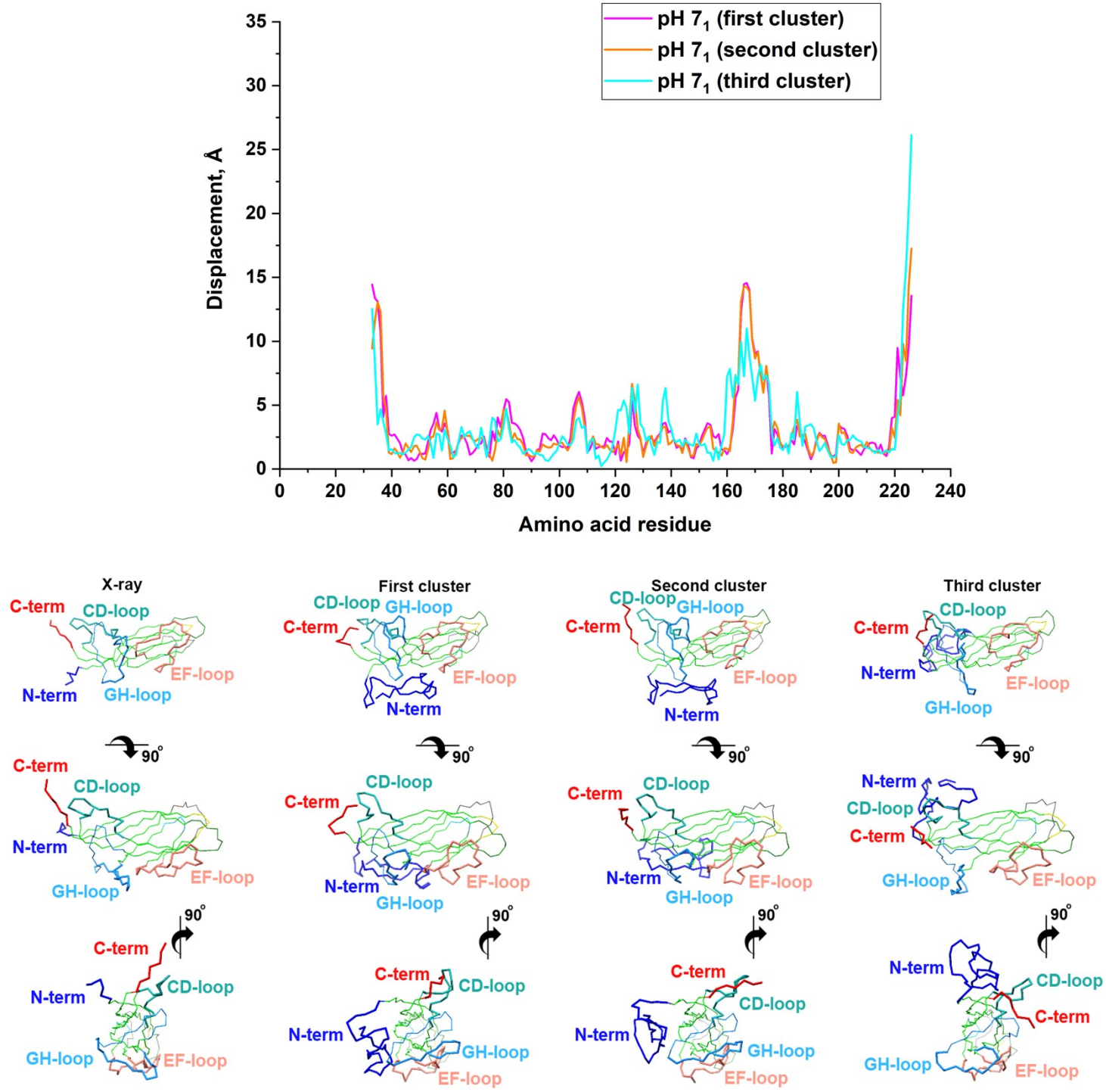
b



Supplementary Figure 18. pH-dependent regions of the CP. (a) Positional displacement of representative structures from first three populated clusters at pH 3 compared to the X-ray structure. (b) Cα trace of representative structures from the first three clusters at pH 3

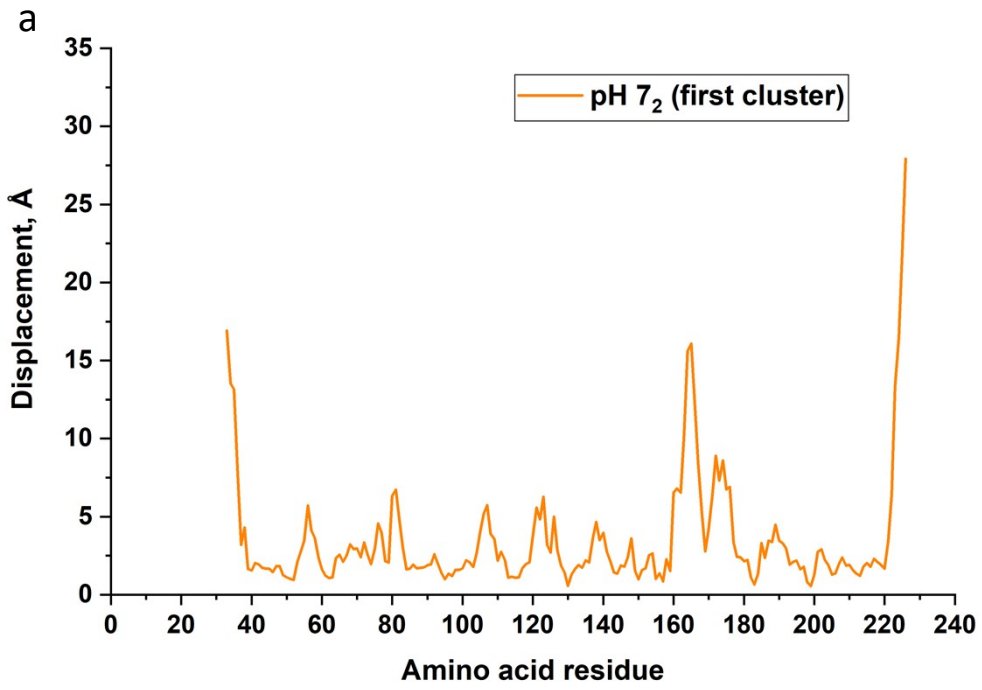


Supplementary Figure 19. pH-dependent regions of the CP. (a) Positional displacement of representative structures from first three populated clusters at pH 5 compared to the X-ray structure. (b) Cα trace of representative structures from the first three clusters at pH 5



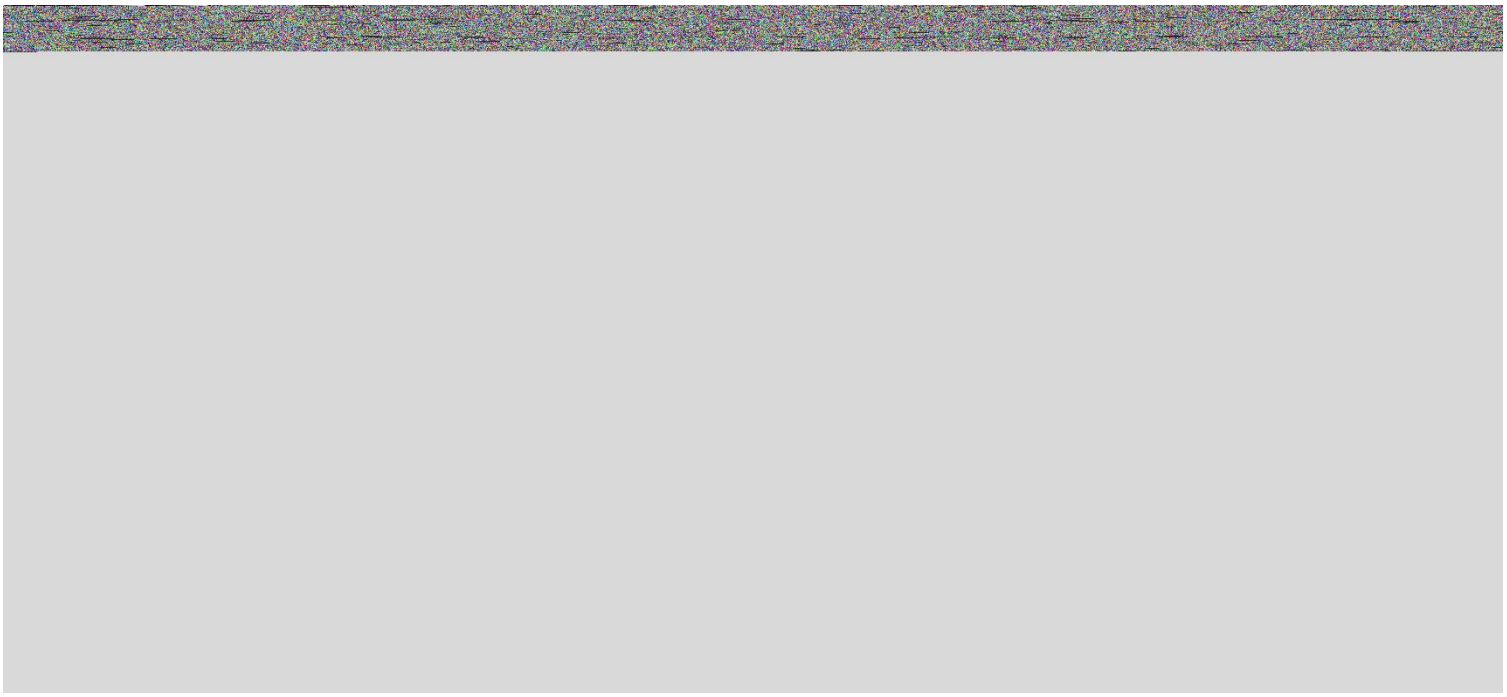
Supplementary Figure 20. pH-dependent regions of the CP. (a) Positional displacement of representative structures from first three populated clusters at pH 7_1 compared to the X-ray structure. (b) Cα trace of representative structures from the first three clusters at pH 7_1

a

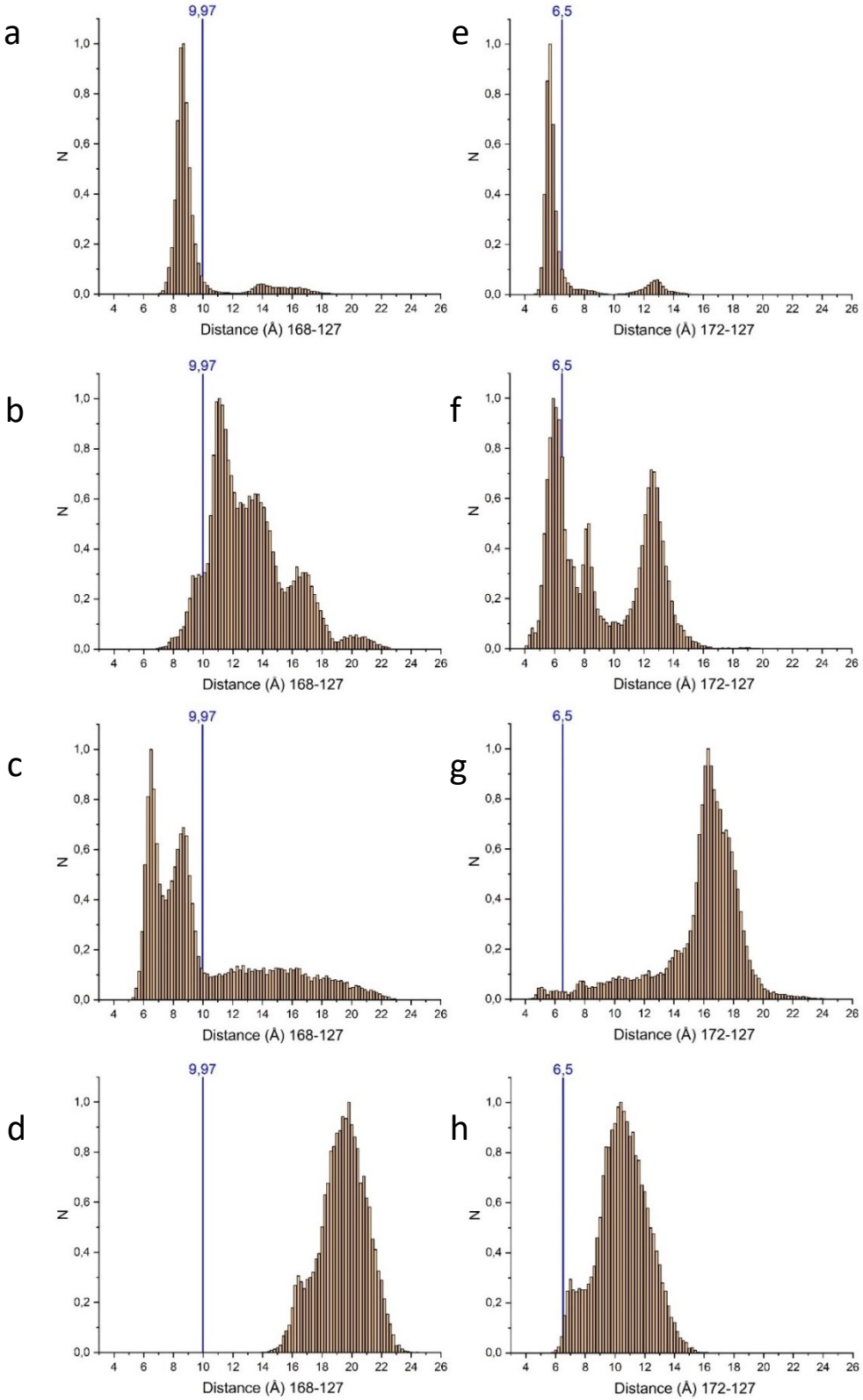


b

b



Supplementary Figure 21. pH-dependent regions of the CP. (a) Positional displacement of representative structures from first three populated clusters at pH 7_2 compared to the X-ray structure. (b) Cα trace of representative structures from the first three clusters at pH 7_2



Supplementary Figure 22. Fractions of extracted structures distributed according to the distances between residues 168 and 127 a,b,c,d and residues 172 and 127 e,f,g,h; at pH 3 – a,e; pH5 – b,f; pH 7 (first minimum) – c,g; pH7 (2 minimum) – d,h. Blue vertical line represents distances between the residues in the capsid structure.